



Data Integration Strategies & Implementation(DISI) Overview

PT-Level Data
Repository Workshop
Session 1
April 22, 2022

Technical Assistance Platform (TAP)

- **Overall Goals**

- Improve the effectiveness and efficiency of HIS, while ensuring security and sustainability
- Improve individual level data collection and management, enabling case surveillance and timely reporting
- Foster better ways of designing, building, and supporting health information systems

TAP Strategy

To enhance existing systems; develop generic, reusable standards-based, interoperable, secure products

Develop Once

Adapt & Reuse

Support countries to configure, pilot, and implement products

Scale, Sustain, & Share

TAP Focus Area Goals

OpenMRS HIV Reference Implementation (OHRI)

Support the delivery of high quality HIV services, incorporating programmatic, technological, standardization, and data management best practices

Data Integration Strategies & Implementation (DISI)

Integrate data from different sources and levels to generate value for stakeholders

Overall HIS Support (OHISS)

Enable policy, governance, capacity, and standards for HIS, health information exchange (HIE), and human resources to support digital health

TAP Partners' Consortiums



University of California
San Francisco

Data Integration Strategies & Implementation (DISI)

- Integrate data from different sources and levels to generate value for stakeholders
 - Process for bringing data together- datasets, data standardization, data flow, data exchange, record linkage, security and procedures
 - Data sources- Information Systems, jurisdictions, data sharing
 - Benefits resulting from data integration – improved data quality, analytics and visualization, data use
 - The people- patients, program users, ministries of health, agencies ...
- Outcome
 - Create comprehensive view on patients, health programs or epidemics/outbreaks that enables timely decisions on:
 - patient care and service delivery, program monitoring and planning and response to disease outbreaks and epidemics
 - Documentation of DISI processes and artifacts to facilitate adoption by countries

What is expected from DISI?

Practical guidance/ recommendations

for integrating health data from different sources and jurisdictions while ensuring availability of quality data in a manner that ensures privacy, security and confidentiality for disease, case and program management to enhance decision making

Software applications/platforms

with the relevant documentation that enable:

- Customization and deployment
- Data standardization
- Patient matching and record linkage
- Data integration
- Data use

Country implementation

Leverage DISI artifacts and/or processes to implement or improve data integration for their use case

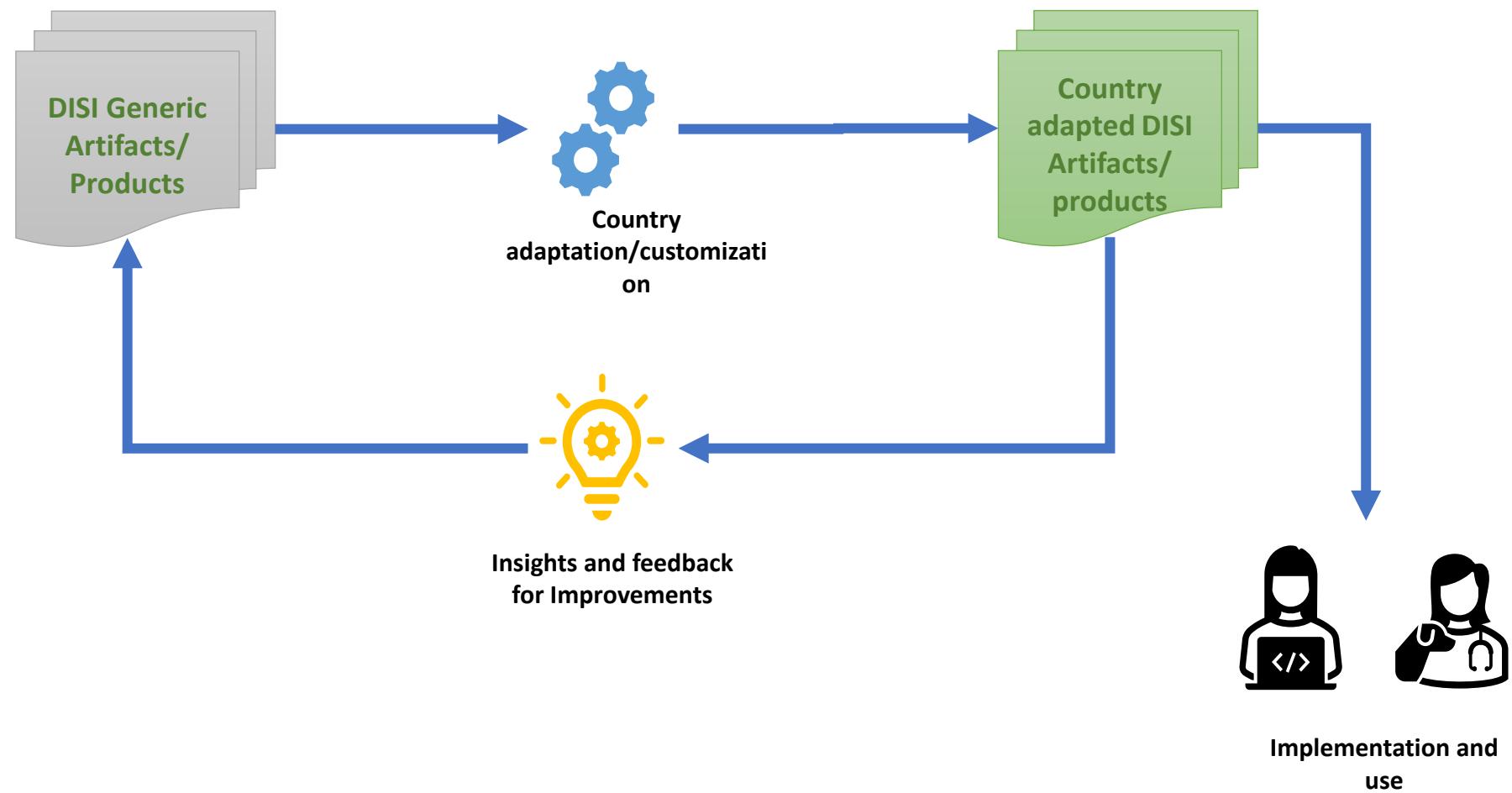
Adopt-adapt-improve approach

Process to monitor and measure

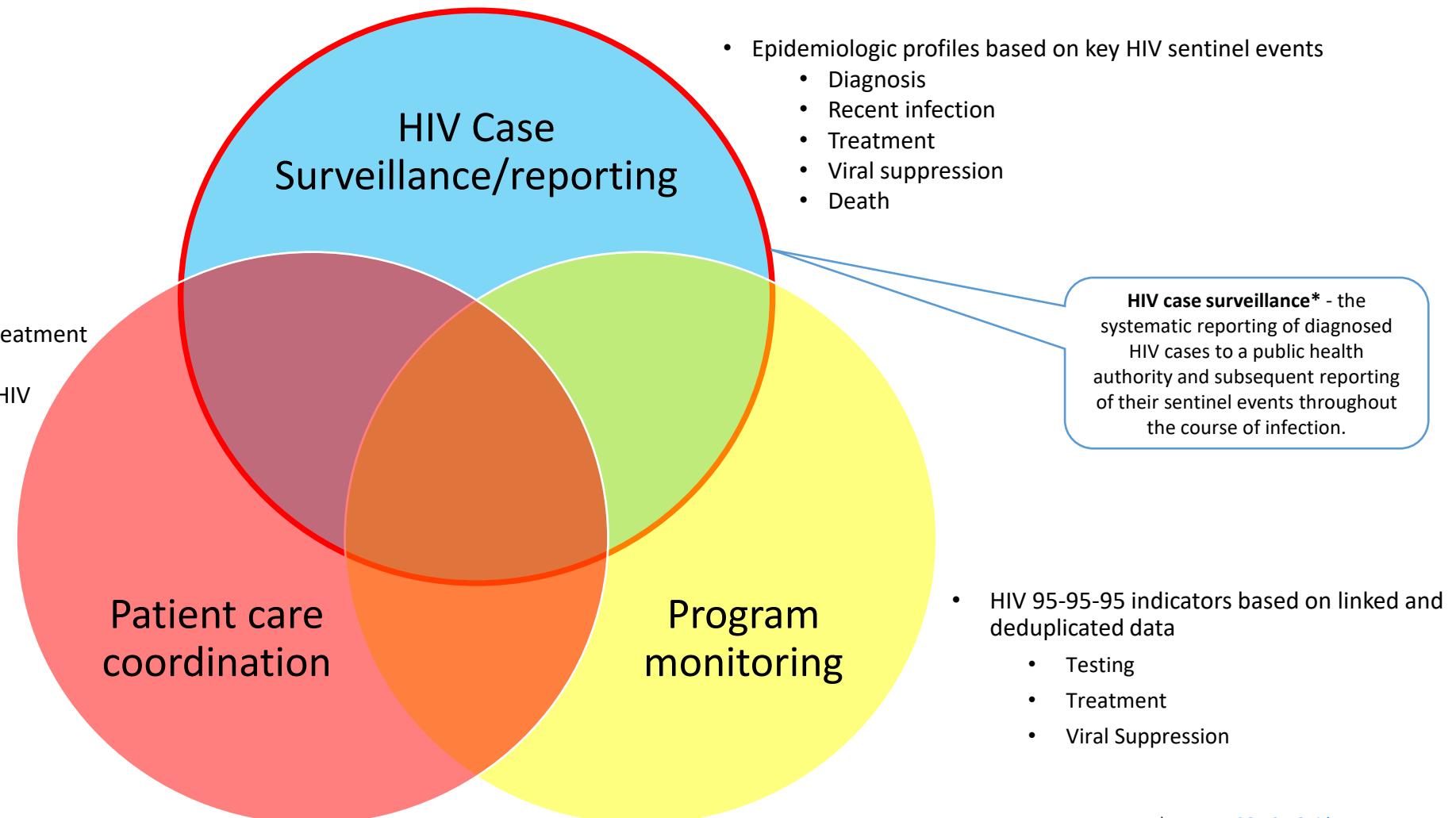
potential benefits and risks of implementing DISI

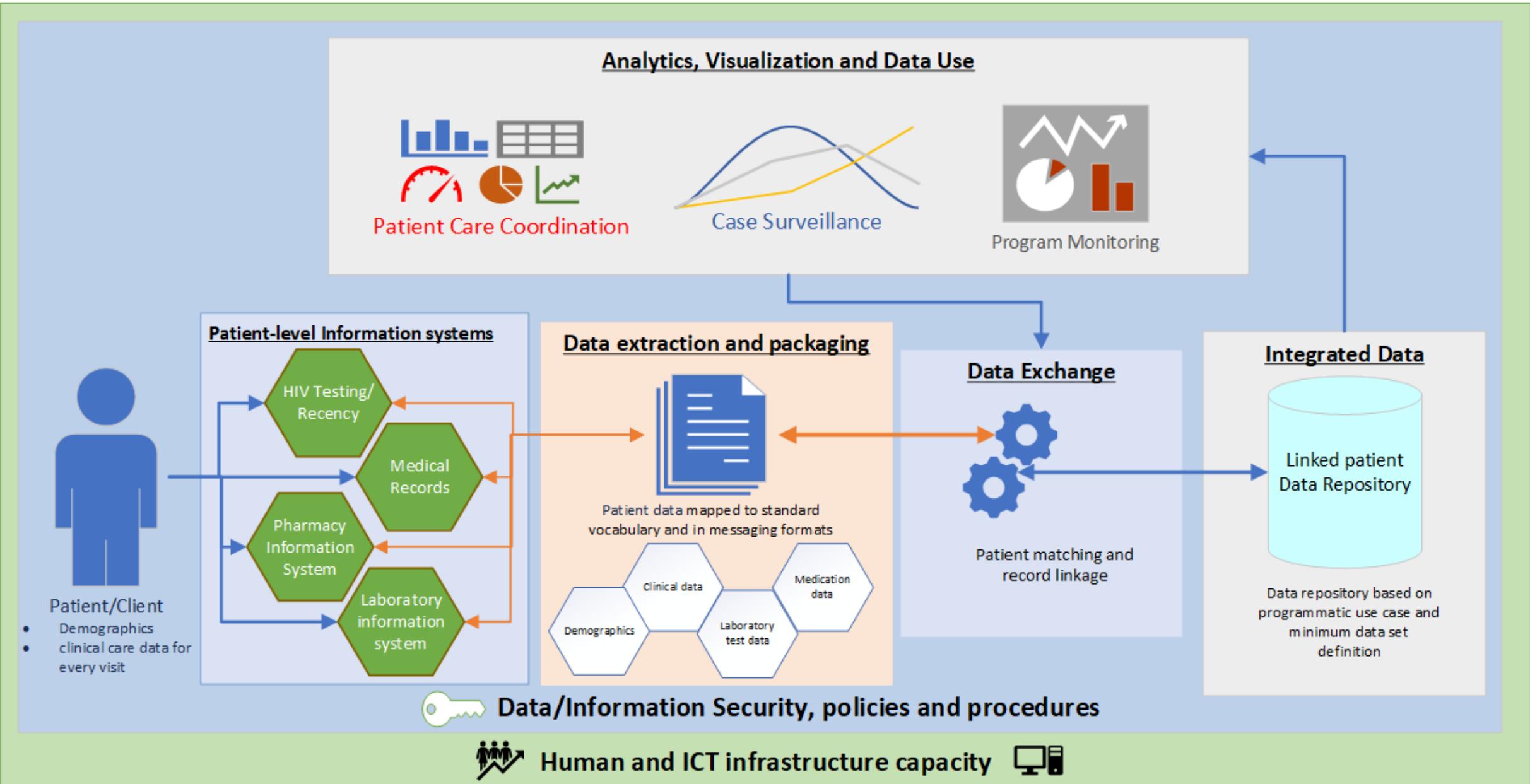
- Metrics for projected efficiencies

Adapt, Adopt and Implement approach



DISI Programmatic Use cases







- Minimum dataset for HIV use case**
 - Mapped to standard vocabulary such as ICD, SNOMED and LOINC
- Application software and platforms**
 - Source applications
 - Interoperability Software platform
- Data exchange workflow and process**
 - Messaging and content format

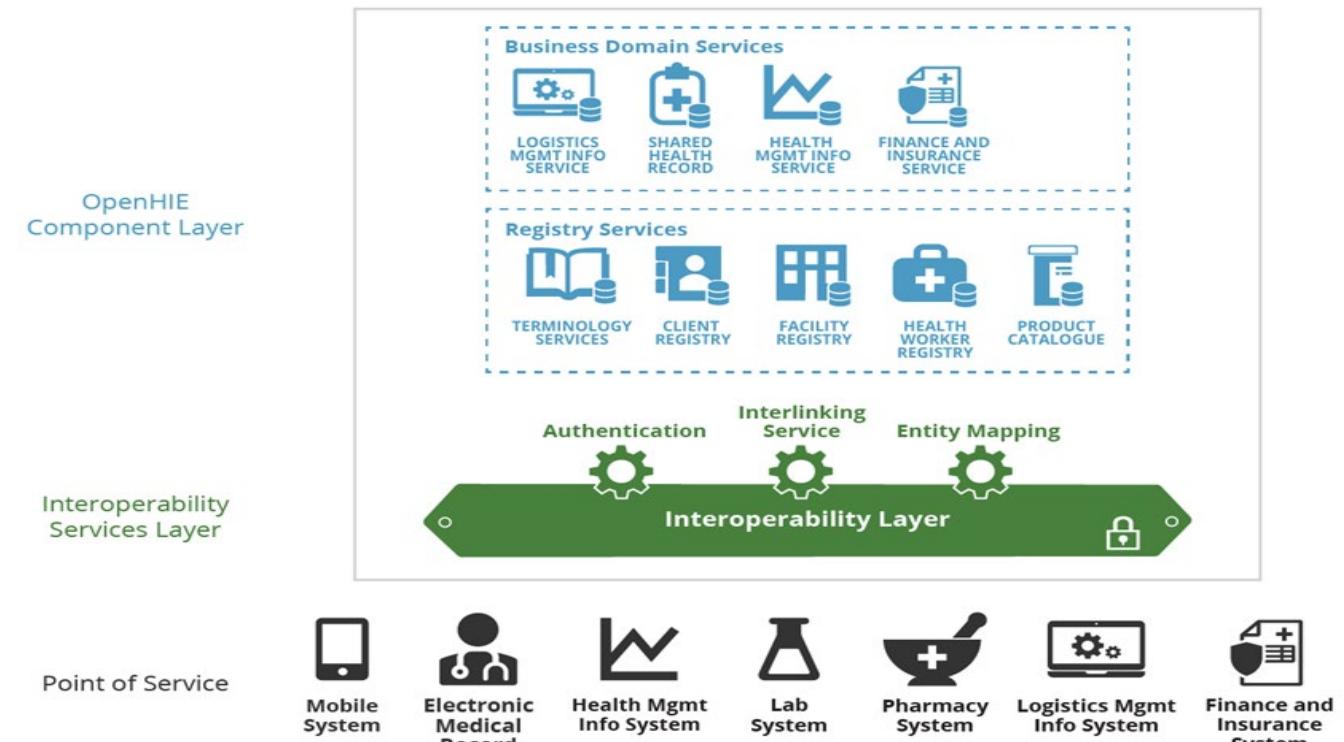
- Privacy security and confidentiality of patient data shared and integrated**
 - Access and use of personally identifiable and sensitive information
- Change management procedures**
 - Managing changes for applications
 - Updates on data exchange processes and workflows

- Training requirements for different aspects of DISI**
 - Technical users
 - Program users
 - Decision makers
- Capacity building/training based on DISI content areas**
- Knowledge, skills, capacity needed to implement, support and maintain**

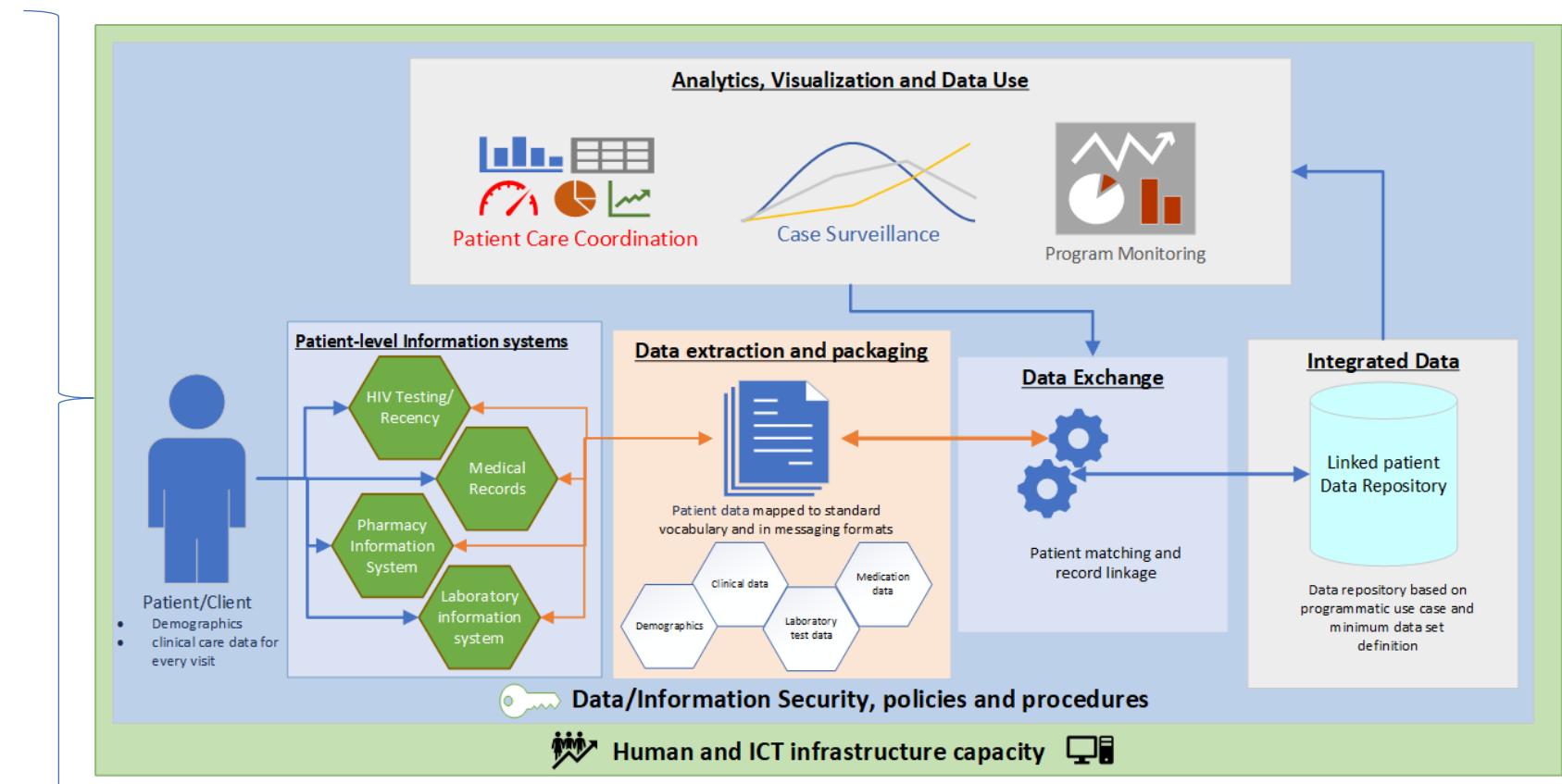
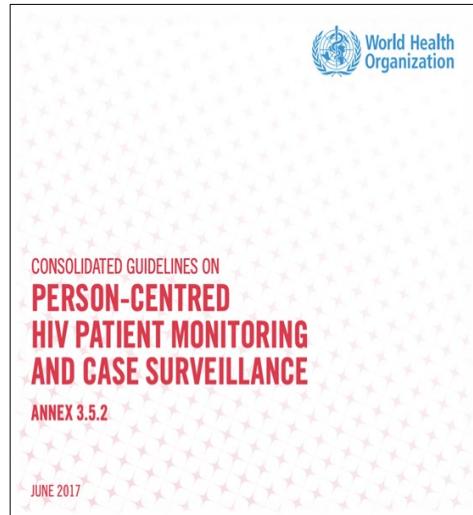
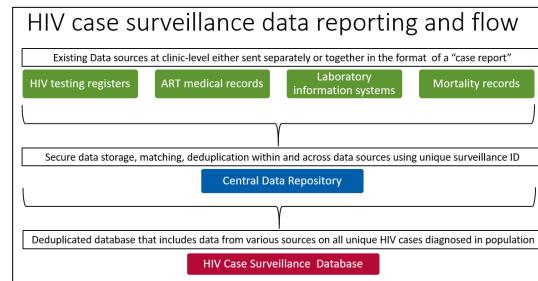
Having client-level HIS infrastructure and health information exchange systems with the ability to uniquely, securely, and confidentially link individual-level patient data to track sentinel events across points of testing, care, clinical monitoring, and treatment sites and deduplicate data is critical to support data collection for Case Surveillance.

Health Information exchange

- Enables data exchange between two or more electronic health information systems.
- Provides data routing functions between information systems
- Facilitates entity matching
- Helps maintain
 - Data integrity – validate data being exchanged
 - Privacy, security and confidentiality – provides centralized authentication, authorization and audit logging services.



HIV CBS implementation - DISI approach



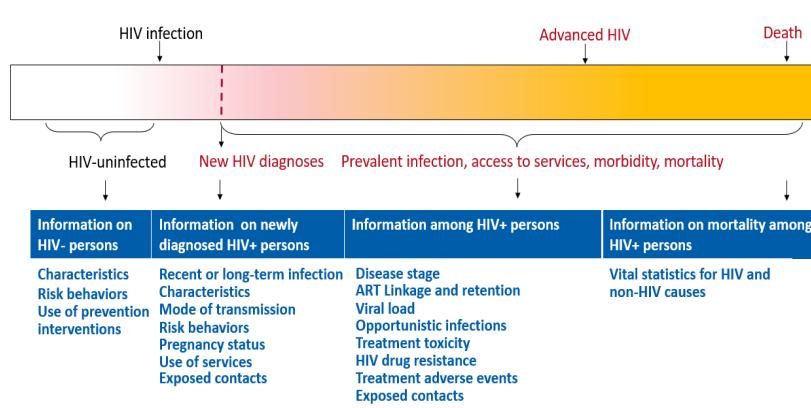
- The WHO has published detailed CBS guidelines. Available via https://www.who.int/hiv/pub/guidelines/WHO_Consolidated_Guidelines_Annexes_3.5.2.pdf?ua=1

HIV CBS Key Components

- Minimum Dataset
- Types of Data and Sources Informing HIV CBS
- Patient Matching and Record Linkage
 - Master Patient Index / Client Registry
 - Deduplication and Linkage
- Health Information exchange (Interoperability Layer)
- **Data Repository**
- Analytics & Visualization Tools

Minimum Dataset

- Specify which data elements (also referred to as variables, indicators, data points or concepts) need to be collected as part of each sentinel event.
- This process identifies WHO, WHAT, WHEN and WHERE



- Data elements mapped to international and local data (vocabulary) standards such as SNOMED, ICD, LOINC, CIEL, RxNORM.**
- Mapping help ensure that data is collected consistently in different information systems and maintains meaning when linked/integrated.**

Sentinel Event	Data Element	Occurrence	Data Type	Where Captured	Data Source	Possible Values
Client registration	SubCounty/District/Arrondissement where client registered	Single	Text	Source	OHRI	
Client registration	Ward/Division/Commune where client currently resides	Single	Text	Source	OHRI	
Client registration	Village/Communal Section where client currently resides	Single	Text	Source	OHRI	
Client registration	Client telephone number	Single	Int	Source	OHRI	
Client registration	Client National Identification	Single	Text	Source	OHRI	
Client registration	Client guardian name	Single	Text	Source	OHRI	
HIV Diagnosis	HIV positive diagnosis date	Single	Date	Source	OHRI	
HIV Diagnosis	HIV positive diagnosis facility site code	Single	Text	Source	OHRI	
HIV Diagnosis	HIV positive diagnosis facility name	Single	Text	Source	OHRI	
HIV Diagnosis	Client HIV positive testing unique identifier	Single	Int	Source	OHRI	
HIV Diagnosis	Recency of HIV infection test conducted	Single	Boolean	Multiple	OHRI,REF Lab	Yes No
HIV Diagnosis	Date of HIV recency test	Single	Date	Multiple	OHRI,REF Lab	
HIV Diagnosis	Actual recency test result	Single	Text	Multiple	OHRI,REF Lab	
HIV Diagnosis	Client HIV infection recency status	Single	Text	Derived	OHRI,REF Lab	Recent Not Recent
Entry to Care	Date client enrolled to care	Single	Date	Source	OHRI	
Entry to Care	Client unique ID assigned at enrollment	Single	Text	Source	OHRI	
Entry to Care	Enrolling facility site code	Single	Text	Source	OHRI	
Entry to Care	Enrolling facility site name	Single	Text	Source	OHRI	
Entry to Care	Date of first clinical visit	Single	Date	Source	OHRI	
ART Initiation	Date client initiated on ART	Single	Date	Source	OHRI	
ART Initiation	ART regimen line client initiated on	Single	Text	Source	OHRI	First Line Second Line Third Line
ART Initiation	ART regimen client initiated on	Single	Text	Source	OHRI	TDF/3TC/DTG TDF/3TC/EFV....
First/Baseline CD4	Client baseline CD4 sample collection date	Single	Date	Source	OHRI	
First/Baseline CD4	Client baseline CD4 result	Single	Int	Source	OHRI	
First/Baseline CD4	Client baseline CD4 percentage	Single	Int	Source	OHRI	

Types of Data and Sources Informing HIV CBS



- Data for HIV case-based surveillance is sourced from multiple patient service delivery points to make longitudinal patient record necessary to create a comprehensive view of the disease progression from diagnosis to outcome

- **Types of datasets.**

- HIV Testing Services (HTS) Data
- Care and Treatment Data
- Viral Load Data
- Population Based Statistics
- Vital Statistics/Mortality

- **Common data sources include:**

- EMR Systems
- Community Testing Applications
- Paper Based Records
- Pharmacy Information Systems
- Lab Information Systems
- Paper Based Viral Load Patient Records
- Civil registration of vital events

Patient Matching and Record Linkage

- Patient Matching
 - Identification and linking of one patient's data within and across health systems
 - Typically accomplished by linking multiple demographic data elements such as name, birth date, phone number, and address.
 - Different approaches such as deterministic, probabilistic and machine learning can be used for matching
- Deduplication and Linkage
 - Eliminate duplicate or redundant data within a health information system
 - Find data entries that correspond to the same entity (typically a patient) in multiple sources and logically merge them to a single comprehensive health record.
- Master Patient Index / Client Registry
 - An electronic database containing patient demographic information within a certain jurisdiction
 - typically, a health facility, and extending to national or subnational data repositories, depending on type of implementation.
 - MPI/CR application analyzes the demographic information to:
 - identify matches from different sources,
 - merge patient demographic information for accurately matched patients
 - assign each patient a unique identifier that can be used to link patient records from the disparate systems.

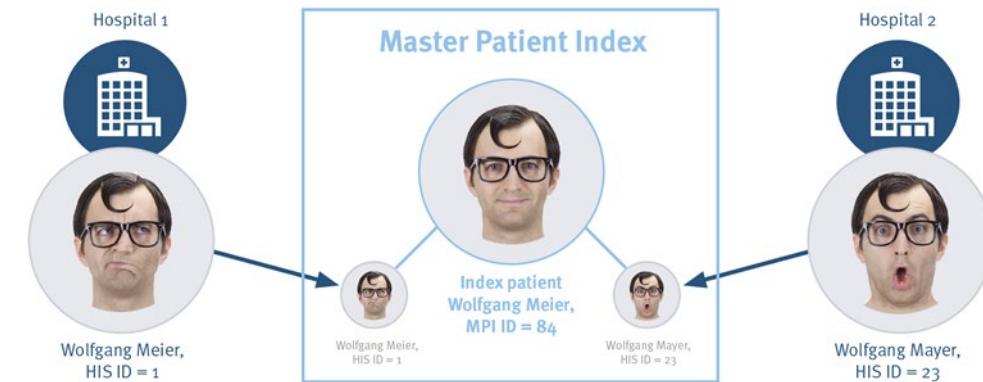
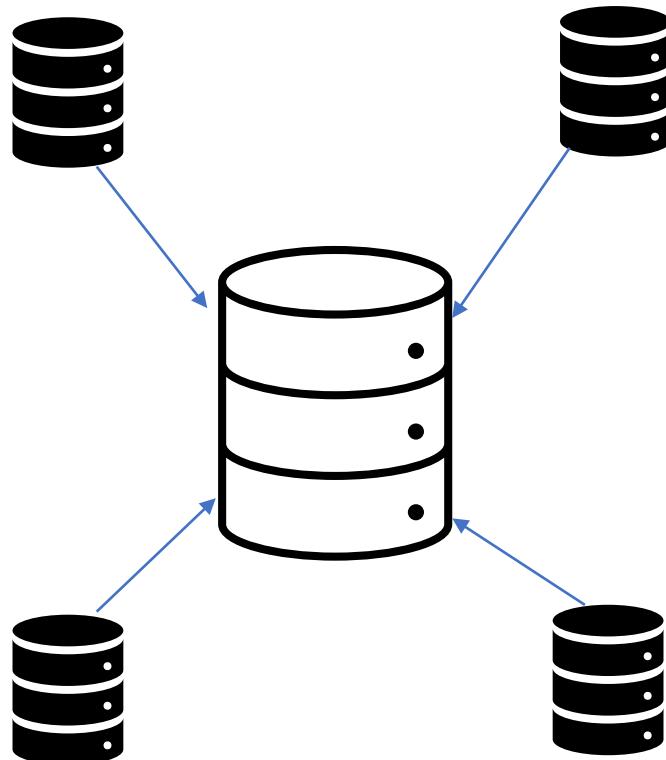


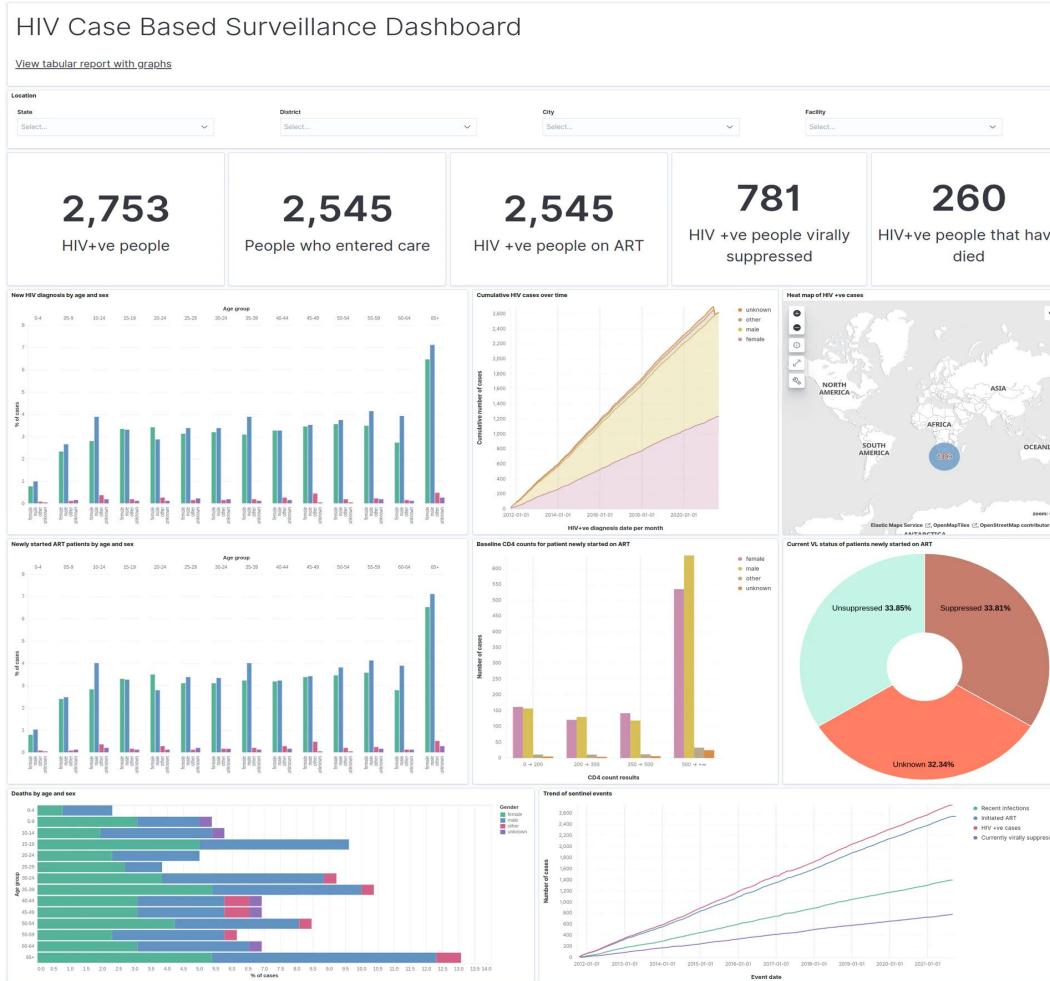
Image taken from <https://icw-global.com/product/master-patient-index/>

Data Repository

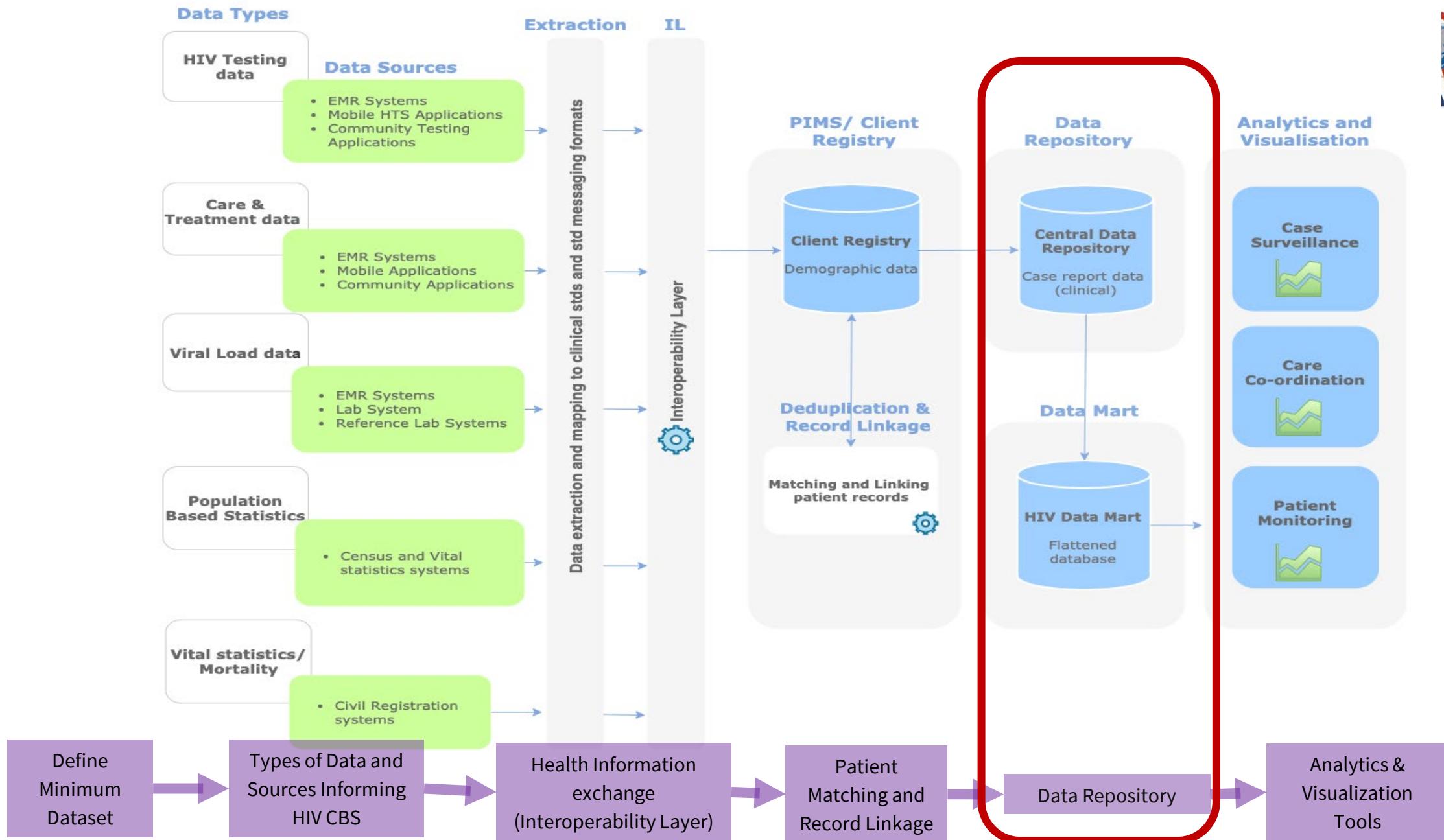


- Manages and stores linked patient records from different sources
 - For HIV case-based surveillance, the data repository would have data on the sentinel events from HIV positive diagnosis to death thus informing HIV epidemiological analysis.
 - Program/disease dataset should inform how the data is organized in the data repository
- Data used for analytics and visualization and can be used by various stakeholders
- Data can be anonymized to safeguard patient privacy and confidentiality while allowing simple and complex data analysis to be conducted
 - Using a client registry that contains patient demographic information can help with anonymization

Analytics & Visualization Tools



- Provides capabilities to process the data into information that can be consumed by various stakeholders
 - Dashboards provide visual representation of information with capability for simple analysis including filtering and grouping.
 - Standard and Ad Hoc reports (usually at National, Sub-National or program levels) can also be generated as end products of a case-based surveillance system
- For HIV case-based surveillance, stakeholders are interested in viewing trends and outcomes typically by demographics or region, all geared towards informing the HIV epidemic.



Discussion, Q&A

The development of HIS products for global use is supported by the U.S. President's Emergency Plan for AIDS Relief (PEPFAR) through the U.S. Centers for Disease Control (CDC) TAP central mechanism under the terms of a cooperative agreement. These products are solely the responsibility of the funding recipients and do not necessarily reflect the views of the United States Government.

